

THE FIDELITY OF THE FOSSIL RECORD: THE IMPROBABILITY OF PRESERVATION

by C. R. C. PAUL

Department of Earth Sciences, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, UK; e-mail: glrcrp@bristol.ac.uk

Typescript received 29 April 2008; accepted in revised form 7 July 2008

Abstract: The fidelity of the fossil record reflects how accurately it preserves the history of life. Since Darwin's time any mismatch between our theories and the fossil record has been attributed to the imperfections of the record. For over a century scarcity of gradual evolutionary trends was explained in this way until the punctuated equilibrium model was proposed. A null hypothesis that all morphological patterns in the fossil record are unbiased random walks can be rejected because it predicts far more apparent trends than exist. Current best estimates suggest that trends occur in at most 5% of characters.

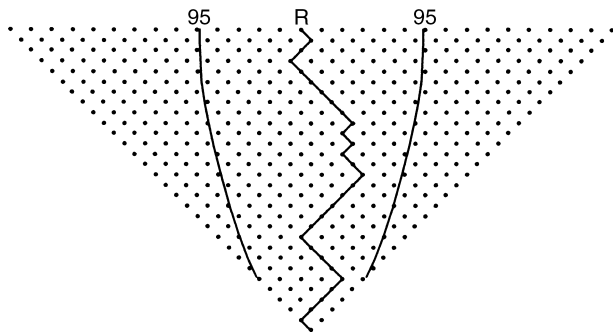
When an organism dies either it becomes fossilized or it doesn't. To be confident a species has not been preserved the probability against preservation must be significantly larger than the total number of individuals of that species that ever existed. For skeletized species preservation was the norm not the exception. Nevertheless, fossils must then avoid subsequent destruction and be discovered to be useful.

Key words: fidelity of the fossil record, evolutionary patterns, preservation probability.

THE FIDELITY of the fossil record reflects how closely it preserves the true history of life on Earth. Clearly there are biases, for example, in preservation potential. Soft-bodied organisms are much less likely to be preserved than those with durable skeletons. The question is, however, do these biases affect the fossil record so seriously that it preserves an unreliable account of the history of life? I believe the answer is 'no', but clearly others disagree, either explicitly (e.g. Patterson 1981) or implicitly by assuming that all discrepancies are due to the incompleteness of the fossil record (e.g. the relative completeness index of Hitchin and Benton 1997). A clear example was the belief that the fossil record should contain abundant evidence of gradual evolutionary (i.e. morphological) change, which we all accepted (myself included) until Eldredge and Gould (1972) proposed the alternative punctuated equilibrium model. The idea of gradual morphological change dates back to Darwin (1859), who devoted two chapters of 'The Origin of Species' to geological evidence, which were largely an attempt to explain why the fossil record did not contain the evidence of gradual change that Darwin expected to see. Durham (1967) repeated Darwin's arguments. He wrote, for example, 'When our knowledge of the fossil record is adequate, the convenient morphological breaks (due to local gaps in the record or to incomplete examination of the record) now used to separate sequential species will no longer

exist, and that our species boundaries will be arbitrary points in a continuum.' (Durham 1967, p. 560).

Eldredge and Gould (1972) presented an alternative view that most morphological change occurs at speciation events, with long periods of stasis (i.e. no morphological change) in between. Although the idea generated considerable debate at the time, it is now widely accepted. Paul (1999) argued that the dominance of stasis in the fossil record could be demonstrated by erecting a null hypothesis that all morphological patterns are unbiased random walks (Paul 1999; Text-fig. 1). Apparent evolutionary trends become random walks that just happen to be unidirectional and lie outside the 90 or 95% confidence intervals (CI). However, with a large sample such as the entire fossil record, the null hypothesis predicts that 5% of all random walks will lie outside the 95% CI. Thus apparent evolutionary trends would be abundant, especially as they usually involve characters not species. If the average fossil had 20 morphological characters, there would be as many apparent trends as there are fossil species. Thus the null hypothesis can be rejected, not because we know some examples of apparently genuine evolutionary trends, but because there are far too few of them. This is particularly true because for over 100 years palaeontologists sought examples of evolutionary trends in the fossil record and yet they remain stubbornly rare. In contrast, no one reported examples of stasis during this interval. We did



TEXT-FIG. 1. Unbiased random walk with 95% confidence intervals (CI) (95). Dots represent every possible position in a 30-step random walk. R represents a single random walk produced by tossing a coin and moving left one step for heads and right for tails. Random walks that lie outside the 95% CI would appear to be unidirectional evolutionary trends. However, with a million trials one would expect 50 000 random walks to lie outside these limits. With a null hypothesis that all morphological patterns in the fossil record were random walks, apparent trends would be abundant. Clearly this is not true and the null hypothesis can be rejected.

not even have a name for lack of morphological change – it was regarded as lack of information. Even classic examples of supposed evolutionary trends, such as *Micraster* (Rowe, 1899), *Zaphrentis* (Carruthers, 1910) and *Gryphaea* (Trueman, 1922), do not bear close scrutiny. There are either gaps in their fossil records (*Zaphrentis*, *Gryphaea*) or the trends occur in a selected interval of their range (*Micraster*). The last is known from at least the Lower Cenomanian (Smith *et al.* 1988) into the Tertiary (Smith and Jeffrey 2000), a minimum period of 32 Myr. Rowe's classic trends occur in the Turonian to Lower Santonian, an interval of 8 Myr. What was happening to *Micraster* during the rest of its evolutionary history? Did Rowe just select part of a random walk that happened to be unidirectional and resembled a trend?

Recently, Hunt (2006) has developed a maximum likelihood method to distinguish objectively between trends, unbiased random walks and stasis. His method has the advantage that it identifies which of the three hypotheses best fits the observed fossil data. It is not necessary to reject one hypothesis before accepting another. My own arguments (Paul 1999) boil down to the following: trends are exceptionally rare, the null hypothesis of unbiased random walks can be rejected, therefore stasis dominates by default. Hunt's methods are robust. Simulations showed that even very low levels of sampling (0.1%) did not significantly bias the results. This is very important as regards the fidelity of the fossil record. Subsequently, Hunt (2007) analysed the frequency of the three patterns in a large sample (251 characters in 51 taxa), covering benthonic and planktonic microfossils and macrofossils (mammals, fish and molluscs), as well as size, shape and

other characters. He found that in only 13 characters (5.2%) was directional change (trends) best supported, whereas unbiased random walks and stasis were best supported in 123 (49%) and 115 (45.8%) cases. Hunt commented that since there was an historical bias in favour of trends, 5% was probably an overestimate. I would add further that stasis is probably underestimated because until very recently a researcher who found no change in any of the characters analysed would be unlikely to publish the results. There is only one such example in Hunt's data set. Another reason for the different estimates of stasis is that in Hunt's maximum likelihood method, the best model (whichever one) accounted for 77% of available likelihood (Akaike weights), the other two accounting for 20% and 3%, respectively. In only 20% of characters could both suboptimal models be rejected. A second interesting feature to emerge is that stasis best fits size characters far less frequently than shape characters. The reverse is true for unbiased random walks. Other factors, including environment (planktonic vs benthonic), taxonomic group, and duration of the evolutionary sequence had relatively minor effects.

In summary, the search for gradual change in the fossil record is a cautionary tale. In over a century the very rare examples that were found were accepted as evidence of a general pattern in the fossil record, whereas the more abundant patterns, unbiased random walks and stasis, were ignored. The analyses of Paul (1999) and Hunt (2006, 2007) are very different, yet both agree that trends are rare in the fossil record. With the benefit of hindsight it seems amazing that it took so long for us to recognize that the vast majority of fossil species do not change significantly throughout their stratigraphical ranges. Currently I am apprehensive that those palaeontologists who refuse to accept that patterns of occurrence in the fossil record can be used to test the predictions of character analyses are in danger of making the same mistake. We ignore the fossil record at our peril. This note treats perhaps the most basic aspect of possible imperfections of the fossil record by considering the probability (or improbability) of preservation. This aspect of the fossil record has received remarkably little attention.

THE IMPROBABILITY OF PRESERVATION

When any organism dies there are only two possible outcomes as far as the fossil record is concerned: either it becomes fossilized or it does not. The probability of preservation (p) will be small (say 0.0001, i.e. 1 in 10 000) and the chances against preservation correspondingly large ($q = 0.9999$). $p + q = 1$, because there are only two possible outcomes. With repeated trials the overall probability

of preservation (P) increases as $p = (1 - q^n)$, where n is the number of trials. This applies to any organism, so for each species n = the total number of individuals of the species that ever existed (N) because every individual could potentially become a fossil. Thus to be confident that a species has never been fossilized the preservation probability (p) has to be so small that if expressed as one in n , n must be significantly larger than N . Given this relationship, it is worth considering possible values of N .

Imagine an endangered mammal species of which on average 100 individuals die each year. If only teeth are preservable and it had 32, this gives 3200 potential fossils each year. Suppose its duration (the total period of its existence on Earth; Foote and Raup 1996) is one million years. This gives 3.2 billion potential fossils (i.e., $N = 3.2 \times 10^9$). If its bones could be preserved, N increases by another order of magnitude. This is for a species with a small standing population and a relatively short duration. Nevertheless, snails or cephalopods, with one shell per individual that could become fossilized, are perhaps more typical organisms. However, the brackish water hydrobiid snail, *Peringia ulvae* (Pennant), occurs along the Atlantic coasts of Europe and Africa, from northern Norway to Senegal and throughout the Mediterranean (Hayward *et al.* 1996). It has been recorded at densities of 32 000/m² (Bouchet *et al.* 1979). It was originally described from the Welsh shore of the Dee Estuary in Britain, and in the 1980s and 1990s was abundant on the opposite shore between West Kirby and Red Rocks, Hoylake, at the northwest tip of the Wirral Peninsula. There are c. 6 km² of sand banks in that immediate area. The entire Dee Estuary has an estimated 75 km² of exposed sand banks at low tide. Not all of this area is suitable for *P. ulvae*, although it has a very wide salinity tolerance. Nevertheless, even if it averages only one specimen per m² and is confined to the northwest tip of the Wirral that implies a population of six million individuals in the Dee Estuary. Its total annual population over its entire geographic range is unimaginably large and this must be multiplied by its duration to derive N . Furthermore, *P. ulvae* is a relatively restricted snail because it requires reduced salinity in estuarine conditions to survive. A similar-sized marine snail with a similar geographic range would have a significantly larger N . Not surprisingly, *P. ulvae* is known as a Pleistocene fossil (Ellis 1969).

Foote and Miller (2007, p. 21) list estimates of the probability of preservation of genera per unit time (circa 5 Myr) for major taxa (from the data of Foote and Sepkoski 1999). Values vary from 0.05 for polychaetes to 0.8–0.9 for cephalopods. Interestingly, the stratigraphically useful graptolites with organic skeletons have higher values (0.65–0.9) than bivalves (0.45–0.5) or gastropods (0.4–0.55) with carbonate skeletons. Nevertheless, the major groups of skeletized organisms (ostracods, trilo-

bites, bryozoans, brachiopods, echinoids, bivalves, gastropods, cephalopods, conodonts) have values at or above 0.5. Crinoids (0.4) and asterozoans (0.25) plus cartilaginous and bony fish (0.1–0.15, 0.15–0.3) have low values, perhaps reflecting their multi-element skeletons that disintegrate soon after death.

The idea that preservation is common may be looked at in reverse. Clearly many organisms did get preserved and sometimes in large numbers, suggesting that for many species preservation probability was sufficiently large. For example, Newell (1959) recorded approximately 1×10^9 diatoms from a 1 cc sample of diatomaceous earth and reported that the South African palaeontologist Robert Broom had estimated there were 8×10^{11} vertebrate fossils preserved in the Karoo Formation. In these cases n was presumably relatively small compared to N . Using a different approach, Valentine (1989) showed that 77% of Recent mollusc species living in shallow seas of the Californian Province were found as fossils in the Californian Pleistocene. He estimated that with further sampling as much as 85% of living species might be found.

Foote and Miller (2007, p. 21) also estimated global proportions of living taxa known as fossils for the same major groups as they estimated preservation probability. At family level, only corals, polychaetes, malacostracan crustaceans, and cephalopods have proportions below 50%. Brachiopods and bivalves have 77 and 76% of genera known as fossils, respectively.

In summary, it is impossible to escape the conclusion that most skeletized organisms had a total population that was so large that they were fossilized. This conclusion does not mean that all skeletized taxa have equally high preservation probabilities. Organisms living in areas of active erosion are much less likely to be preserved than those living in marine basins where sediments are accumulating. So, for example, Tavare *et al.* (2002) estimated that only 5% of primate species have become fossilized, whereas Foote (1997) estimated that c. 65% of Tertiary mammal species had been preserved.

DISCUSSION

As far as I am aware Newell (1959) was the first person to realize in the current context that events that are highly unlikely in a single trial (such as winning the national lottery) become quite common if enough trials are performed (so usually someone wins the lottery each week because so many tickets are sold). The chances of preservation in the fossil record are a similar phenomenon. There are just so many potential fossils each year that whatever the chances against preservation large numbers of fossils still get preserved. Newell assumed a one in a million chance of preservation 'of a really good sample'

and showed that in ten million years there was only a one in a hundred chance that such a sample would not have been preserved (Newell 1959, p. 495). Further he assumed 12 million years as the average duration of a geological epoch, which he regarded as the upper limit for the duration of a fossil species. Thus he was able to conclude that a good sample of a skeletized species would have been preserved before the species became extinct. One might question some of the assumptions made, and values used, by Newell, but I think his fundamental conclusion is correct.

In summary, it would seem that for any organism, even a soft-bodied one, the chances against preservation have to be astronomically high if not a single example ever became fossilized. In many cases this was probably true, but for skeletized organisms it seems likely that preservation was the norm, not the exception. Of course, preservation alone does not imply that the organism is known to science. Subsequent erosion, diagenesis or metamorphism may have destroyed any examples that were preserved. For example, Foote and Miller (2007, p. 24) quote Valentine's (1989) conclusion that 77% of bivalve and gastropod species living in the Californian Province are known as fossils in the Californian Pleistocene, yet they estimate that only 1% of easily preserved species have been fossilized. For an alternative view, see Stanley (2007). Even if the fossils have survived, palaeontologists still have to find, recognize and describe them before they become available for other scientific purposes. Raup (1976) pointed out a strong correlation between known fossil species from each geological period and the area of outcrop or volume of rock for each period, suggesting that our knowledge of the fossil record is biased by available rock outcrop. The link between the rock record and biodiversity has been investigated more recently by Peters (2005), Smith and McGowan (2007) and McGowan and Smith (2008), among others.

Acknowledgements. The manuscript was significantly improved by the critical comments of S. K. Donovan, Nationaal Natuurhistorisch Museum, Leiden and two anonymous reviewers.

REFERENCES

- BOUCHET, P., DANRIGAL, F. and HUYGHENS, C. 1979. *Living seashells. Molluscs of the English Channel and Atlantic Coasts*. Blandford Press, Poole, Dorset, 144 pp.
- CARRUTHERS, R. G. 1910. On the evolution of *Zaphrentis delanoui* in Lower Carboniferous times. *Quarterly Journal of the Geological Society*, **66**, 523–538.
- DARWIN, C. 1859. *The origin of species by means of natural selection or the preservation of favoured races in the struggle for life*. John Murray, London.
- DURHAM, J. W. 1967. The incompleteness of our knowledge of the fossil record. *Journal of Paleontology*, **41**, 559–565.
- ELDRIDGE, N. and GOULD, S. J. 1972. Punctuated equilibria: an alternative to phyletic gradualism. 82–115. In SCHOPF, T. J. M. (ed.). *Models in paleobiology*. Freeman, Cooper, San Francisco.
- ELLIS, A. E. 1969. *British snails. A Guide to the Non-marine Gastropoda of Great Britain and Ireland Pleistocene to Recent*, 2nd edn. Clarendon Press, Oxford, 298 pp.
- FOOTE, M. 1997. Estimating taxonomic durations and preservation probability. *Paleobiology*, **23**, 278–300.
- and MILLER, A. I. 2007. *Principles of paleontology*, 3rd edn. Freeman, New York, XV + 354 pp.
- and RAUP, D. M. 1996. Fossil preservation and the stratigraphic ranges of taxa. *Paleobiology*, **22**, 121–140.
- and SEPKOSKI, J. J. Jr 1999. Absolute measures of the completeness of the fossil record. *Nature*, **398**, 415–417.
- HAYWARD, P., NELSON-SMITH, T. and SHIELDS, C. 1996. *Sea shore of Britain and Europe*. HarperCollins, London, 352 pp.
- HITCHIN, R. and BENTON, M. J. 1997. Congruence between parsimony and stratigraphy: comparisons of three indices. *Paleobiology*, **23**, 20–32.
- HUNT, G. 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology*, **32**, 578–601.
- 2007. The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 18404–18408.
- MCGOWAN, A. J. and SMITH, A. B. 2008. Are global Phanerozoic diversity curves truly global? A study of the relationship between regional rock records and global Phanerozoic marine diversity. *Paleobiology*, **34**, 80–103.
- NEWELL, N. D. 1959. Adequacy of the fossil record. *Journal of Paleontology*, **33**, 488–499.
- PATTERSON, C. 1981. Significance of fossils in determining evolutionary relationships. *Annual Reviews of Ecology and Systematics*, **12**, 195–223.
- PAUL, C. R. C. 1999. Selection and the fossil record. 1–13. In BINTLIFF, J. (ed.). *Structure and contingency. Evolutionary processes in life and human society*. Leicester University Press, London, xxii + 153 pp.
- PETERS, S. 2005. Geologic constraints on the macroevolutionary history of marine animals. *Proceedings of the National Academy of Sciences*, **102**, 12326–12331.
- RAUP, D. M. 1976. Species diversity in the Phanerozoic: an interpretation. *Paleobiology*, **2**, 289–297.
- ROWE, A. W. 1899. An analysis of the genus *Micraster*, as determined by rigid zonal collecting from the zone of *Rhynchonella cuvieri* to that of *Micraster coranguinum*. *Quarterly Journal of the Geological Society*, **55**, 494–547.
- SMITH, A. B. and JEFFREY, C. H. 2000. Maastrichtian and Palaeocene echinoids: a key to world faunas. *Special Papers in Palaeontology*, **63**, 1–406.
- and MCGOWAN, A. J. 2007. The shape of the Palaeozoic marine diversity curve: how much can be predicted from

- the sedimentary rock record of Western Europe? *Palaeontology*, **50**, 765–774.
- PAUL, C. R. C., GALE, A. S. and DONOVAN, S. K. 1988. Cenomanian and Lower Turonian echinoderms from Wilmington, south-east Devon, England. *Bulletin of the British Museum (Natural History) Geology Series*, **42**, 1–245.
- STANLEY, S. M. 2007. An analysis of the history of marine animal diversity. *Paleobiology Memoir*, **4**, (Suppl. vol. 33, no. 4), 1–55.
- TAVARE, S., MARSHALL, C. R., WILL, O., SOLIGO, C. and MARTIN, R. D. 2002. Using the fossil record to estimate the age of the last common ancestor of extant primates. *Nature*, **416**, 726–729.
- TRUEMAN, A. E. 1922. The use of *Gryphaea* in the correlation of the Lower Lias. *Geological Magazine*, **54**, 256–268.
- VALENTINE, J. W. 1989. How good was the fossil record? Clues from the Californian Pleistocene. *Paleobiology*, **15**, 83–94.